Can Risk Aversion Explain Schooling Attainments? Evidence From Italy*

Christian Belzil
Centre de National de Recherche Scientifique,
Institute for the Study of Labor (IZA)
and CIRANO

Marco Leonardi
Universita di Milano and
Institute for the Study of Labor (IZA)

February 17, 2006

Abstract

Using unique Italian panel data, in which individual differences in behavior toward risk are measured from answers to a lottery question, we investigate if (and to what extent) risk aversion can explain differences in schooling attainments. We formulate the schooling decision process as a reduced-form dynamic discrete choice. The model is estimated with a degree of flexibility virtually compatible with semi-parametric likelihood techniques. We analyze how grade transition from one level to the next varies with preference heterogeneity (risk aversion), parental human capital, socioeconomic variables and persistent unobserved (to the econometrician) heterogeneity. We present evidence that schooling attainments decrease with risk aversion, but despite a statistically significant effect, differences in attitudes toward risk account for a modest portion of the probability of entering higher

*We thank David Card and Winfried Koeniger for useful comments.

education. Differences in ability(ies) and in parental human capital are much more important. in the most general version of the model, the likelihood function is the joint probability of schooling attainments, and post-schooling wealth and risk aversion.

JEL Classification: J24.

1 Introduction and Motivation

The connection between individual attitudes toward risk and investment behavior has been widely analyzed in financial economics. This is true both at the theoretical and at the empirical level.[1] Although human capital is undoubtedly the main component of individual assets, the link between risk aversion and human capital accumulation, and in particular schooling, remains largely hypothetical. Most of the work is theoretical and often confined to relatively simple two-period models. In general, the results stress that earnings uncertainty may depress human capital investment.[2]

Empirical work remains scarce and is rather inconclusive. There is one main reason for this. At the empirical level, determining which asset is more risky is a relatively straightforward econometric question. However, quantifying the marginal risk which characterizes the transition from one level of schooling to the next is a more difficult research agenda. Not surprisingly, economists are currently unable to say if (and to what extent) schooling acquisition is a risky investment although the issue is starting to raise a significant level of interest. Moreover, the degree of education selectivity based on individual differences in risk aversion is completely unknown.

In this paper, we investigate whether risk aversion can explain differences in schooling attainments. We ask three simple questions. Does risk aversion increase or decrease investment in higher education? How does the effect of risk aversion compare with the effects of ability and family human capital? How much of the cross-sectional dispersion in schooling attainments is explained by differences in attitudes toward risk.

In order to answer these questions, we take an approach completely different from what is found in the literature. Using unique Italian panel data, in which an individual specific measure of risk aversion is inferred from an

[1] See Kocherlakota (1996) for a comprehensive survey.

[2] This is the case, for instance, in Lehvari and Weiss (1974) and Olson, White and Sheffrin (1979).

answer to a lottery question, we formulate the schooling decision process as a reduced-form dynamic discrete choice problem (using discrete duration model techniques) and we analyze how grade transition from one level to the next varies with measured risk aversion. In particular, we decompose the probability of entering higher education into four groups of variables; preference heterogeneity (risk aversion), persistent unobserved (to the econometrician) ability heterogeneity, parental human capital (parents' education and occupation) and socioeconomic variables (sex, region, age cohort).

Our analysis is based on a sample of Italian individuals. Our methodology therefore relies on the fact that higher education in Italy must be a "reasonably" risky investment. While tuition fees are low in Italy (and typically everywhere in Europe), there is no reason to believe that Italian students face lower psychic costs than do students in other countries.[3] For the sake of comparison, Italian students face a relatively more incomplete capital market than do US students. Borrowing while in school is practically inexistent in Italy.[4] The US, on the other hand, has very high tuition rates but also has substantial student loan and fellowship programs. Interestingly, both Italy and the US are characterized by a relatively high level of inequality. Although cross sectional wage dispersion is higher in the US than in Italy, long run (lifetime) inequality is thought to be higher in Italy and, in particular, among the highly educated.[5] To the extent that the riskiness of the education investment may be at least correlated with the individual's lifetime inequality, these institutional facts seem to indicate that investing in higher education may be as risky in Italy as in the US.

Aside from its direct contribution to the revived debate on the schooling/risk trade-off, this paper also contributes to the already existing literature on the determinants of schooling attainments. As of now, labor economists have paid a particular attention to the importance of parental human cap-

[3] Empirical evidence for the US suggests that differences in psychic costs may be quite important. For instance, the large explanatory power of the individual specific differences in the per-period utility of attending school found in the structural literature is consistent with the existence of strong psychic costs (Keane and Wolpin, 1997, and Belzil and Hansen, 2002). See Heckman, Lochner and Todd (2005) and Belzil (forthcoming) for surveys.

[4] The Italian national statistical office (ISTAT, 2003, Table 1.8) reports that the total number of student loans in Italy in the academic year 1999-2000 was 97.

[5] In a recent paper, Flinn (2002) shows that after taking into account job offer probability while employed and while unemployed and unemployment incidence, lifetime welfare inequality is higher in Italy than in the US. His results are obtained in a search framework with risk-neutral workers.

ital and individual abilities (observed or unobserved). This paper adds a new dimension to the analysis of the determinants of schooling; namely the importance of preference heterogeneity.[6]

In line with most of the theoretical literature, we find that more risk averse individuals tend to get less schooling. However, we find that risk aversion is a far less important determinant of schooling attainment than individual-specific abilities and parental human capital (education and occupation). As an illustration, the range of higher education participation probabilities spanned by the 10th-90th percentile of the risk aversion distribution is ten to twenty times **smaller** than the equivalent range obtained for the unobserved heterogeneity term. More specifically, preference heterogeneity accounts for 10% or less of the cross-sectional dispersion in the probability of entering higher education. Our main findings are quite robust. They are robust to the allowance for alternative measures of risk aversion. They are also invariant to the allowance for a differentiated effect of risk aversion at different grade levels.

The paper is constructed as follows. In Section 2, we present some background material and review the most important literature. In section 3, we discuss the Bank of Italy Survey of Income and Wealth (SHIW) and provide details about the measure of risk aversion used in our analysis. The econometric model is described in Section 4. Section 5 contains a brief overview of the model specifications. In Section 6, we present evidence that schooling attainments decrease with risk aversion and we compare the effects of risk aversion with the effect imputed to unobserved heterogeneity. In the following section, we decompose the total variation in the probability of entering higher education into its main components, preference heterogeneity, ability(ies), parental background, and socio economic attributes. The conclusion is found in Section 8.

2 Background an Relevant Literature

Fundamentally, the marginal risk associated to schooling has two distinct components. One component relates to the human capital accumulation process and is experienced by the individuals at the time schooling decisions

[6] In the structural literature, the term "preference heterogeneity" is often used to refer to differences in taste for schooling and academic abilities (Keane and Wolpin, 1997). In our analysis, these unobserved factors are subsumed in the unobserved heterogeneity term.

are made. The second component relates to post-schooling labor market outcomes and is therefore associated to the (perceived) distribution of random variables which are realized much beyond actual schooling decisions.

With respect to the accumulation process, acquiring schooling should be unambiguously viewed as a risky investment. School (and especially college) attendance requires to sacrifice present consumption and to absorb substantial psychic costs in return for future rewards, but successful grade achievement is rarely a certain outcome. For this reason, the probability of losing the investment paid up front cannot be ignored and may act as a strong disincentive.

At the level of labor market outcomes, the argumentation becomes more complicated. In practice, life cycle earnings are affected by random events such as job offers, layoffs, risk sharing agreements between firms and workers (or unions) and many other events. Occupation choices may also affect earnings volatility. The ex-ante probability distribution of those labor market outcomes may depend on schooling but it is far from clear if accumulated schooling contributes to an increase in earnings dispersion or decreases volatility.[7] On top of this, wages and earnings are typically affected by individual ability heterogeneity as well as by idiosyncratic and business cycle shocks. Separating these components may be particularly difficult (Cunha, Heckman, and Navarro, 2005).

In the long run, labor market productivity and earnings may be affected by structural changes in the economy. Potential technological changes affecting the return to schooling may be viewed as an additional element of risk from the perspective of the student. On the other hand, when schooling is viewed as facilitating adjustment to technological change, this uncertainty may turn out to favor schooling acquisition (i.e.: schooling becomes a form of insurance).[8]

Given this level of complexity, and taking into account both the accumulation process and labor market outcomes, it is difficult to say whether or not individuals perceive schooling acquisition as a truly risky investment. In the earlier literature, a few descriptive analyses of the variability of empirical age/earnings profile have been carried out. However, the notion of variability

[7] For instance, schooling may reduce earnings dispersion by reducing the unemployment incidence or by raising the job offer probabilities (given unemployment) but it may increase wage volatility if more educated workers find jobs in sectors or occupations where wages (or marginal product) is more volatile.

[8] This argument is put forward in Gould, Moav and Weinberg (2001).

is usually an "ex post" notion which may have little to do with "ex ante" risk.[9] Ideally, evaluating the marginal risk would require a statistical analysis of the joint distribution of life cycle wages, unemployment, job offer probabilities and grade completion (or failure) probabilities. In particular, it would also require to disentangle persistent unobserved (from the econometrician perspective) heterogeneity from true dispersion. This would be difficult to achieve and indeed, as of now, such a comprehensive study does not exist.

On top of this, measuring the marginal risk associated to schooling for all relevant labor market outcomes may turn out to be irrelevant if individuals have imperfect information about the law of motion that generates labor market outcomes. If so, individual subjective probabilities may diverge from the Rational Expectation hypothesis and the use of post-schooling panel data on wages and employment outcomes may become irrelevant for the econometrician.[10]

As it stands now, there is no strong empirical evidence on the effect of education on wage/earnings dispersion, but economists are starting to pay more and more attention to the issue. In a recent paper, Palacios-Huerta (2003) presents an empirical comparison of the properties of risk-adjusted rates of return to schooling within an intertemporal model, using mean-variance spanning techniques.[11] In line with the stream of the literature devoted to the increase in wage inequality, many individuals have analyzed the wage dispersion (basically the variance) within education groups in cross-section data rather than in panel data. The cross-section evidence shows that the variance of wages is higher within the educated group (Lemieux , 2005 and Chay and Lee, 2000). In an attempt to separate individual heterogeneity from ex-ante risk, Belzil and Hansen (2004) estimate a dynamic programming model in which the degree of risk aversion can be inferred from schooling decisions but they assume that the attitude toward risk is represented by a parametric (constant relative risk aversion) utility function. They identify the degree of risk aversion from the degree of heteroskedasticity in the idiosyncratic earn-

[9] Mincer (1974) investigates how the variance of earnings differs across schooling levels over the life cycle while Chiswik and Mincer (1972) use age earnings profile to investigate time series changes in income inequality. Kodde (1985) uses the Lehvari and Weiss model as a background for empirical work and tests predictions from the model from data on subjective estimates (self reported) of future earnings.

[10] See Geweke and Keane (2001) for an insightful discussion.

[11] Basically, the mean-variance spanning technique amounts to quantifying the effect of introducing a new asset on the mean-variance of another benchmark asset.

ings shock but assume that all persistent unobserved heterogeneity is in the information set of the agent. As panel data on wages, earnings and schooling do not allow them to identify cross-sectional dispersion in risk aversion, they assume homogeneity of preferences and automatically rule out the possibility that differences in schooling are driven by differences in attitudes toward risk. Finally, Cunha, Heckman and Navarro (2005) develop a statistical method which distinguishes between heterogeneity and risk but also allow for a distinction between ex-ante risk and ex-post dispersion. Their method allows the econometrician to infer the set of variables upon which schooling decisions are based, but disregards heterogeneity in risk aversion.[12]

3 Measuring Risk Aversion: The Bank of Italy Survey of Income and Wealth

We use data from the 1995 wave of the Bank of Italy Survey of Income and Wealth (SHIW). The survey collects information on consumption, income and wealth in addition to several household characteristics for a representative sample of 8,135 Italian households. More importantly, the 1995 survey contains a question on household willingness to pay for a lottery which can be used to build a measure of individual risk attitudes.

In the survey, each head of household is asked to report the maximum price he/she is willing to pay to participate to an hypothetical lottery. The question is worded as follows:

"We would now like to ask you a hypothetical question that we would like you to answer as if the situation was a real one. You are offered the opportunity of acquiring a security permitting you, with the same probability, either to gain a net amount of 10 million lire (roughly 5,000 dollars) or to lose all the capital invested. What is the most you are prepared to pay for this security?"[13]

The respondent can answer in three possible ways: 1) give the maximum

[12] On top of these few papers, a relatively large number of related working papers are being currently circulated. These include Hartog, Van Ophem and Bajdechi (2004), Chen (2003), Harmon, Hogan and Walker (2003) and Davis and Willen (2002).

[13] In other words, the expected value of entering the lottery is $0.5 \cdot (10,000,000 - bet)$. Guiso and Paiella (2004) write that the interviews were conducted by professional interviewers at the respondents' homes and to help the respondent to understand the question the interviewers showed them an illustrative card and were ready to provide explanations.

price he/she is willing to pay, which we denote as *bet*; 2) don't know; 3) don't want to participate. Of the 8,135 heads of household, 3,458 answered they were willing to participate and reported a positive maximum price they were willing to bet (prices equal to zero are not considered a valid response).[14]

at a theoretical level, it is easy to show that there is a one-to-one correspondence between the value attached to the lottery, and the degree of risk aversion. For a given of wealth, w_i, and a potential gain (g_i), the optimal bet, bet_i, must solve the expected utility equation;

$$U_i(w_i) = \frac{1}{2}U_i(w_i + g_i) + \frac{1}{2}U_i(w_i - bet_i) = EU(w_i + R_i) \tag{1}$$

where R_i represents the return (random) of the lottery. Taking a secon-order expansion, and noting that R_i is also the maximum purchase price (bet_i), we get that

$$EU(w_i + bet_i) \approx U_i(w_i) + U_i'(w_i).E(R_i) + \frac{1}{2}U_i^{''}(w_i) \cdot E(R_i)^2 \tag{2}$$

It is therefore possible to express risk aversion (say the Arro-Pratt measure given by $\frac{-U_i^{''}(w_i)}{U_i'(w_i)}$) as a function of the parameters of the lottery and the the value of the bet of each individual. In general, the optimal bet depends on $U_i(.)$ and on consumer endowment (w_i).The valid responses to the question - *bet* - range from 1,000 lire to 100 million lire and constitute our measure of individual risk aversion. Of the 3,288 heads in our final data set (see the sample selection criteria below), 3,131 reported a maximum price *bet* less than 10 million lire which implies that they are risk averse individuals, 117 reported *bet* exactly equal to 10 million lire (i.e. they are risk neutral) and 40 reported *bet* more than 10 million indicating that they are risk lovers. The empirical distribution of *bet* is reported in Table 2. Although the majority of the respondents are risk averse and only 5% of the sample is either risk-neutral or risk-loving, there is a large heterogeneity in the degree of risk

[14] Guiso and Paiella (2005) also explain that the question has a large number of non responses because many respondents may have considered it too difficult. This does not mean that those who responded gave erroneous anwers. However, the literature in experimental economics (Kagel and Roth, 1995) underlines that individuals tend to report lower buying than selling prices when asked to price hypothetical lotteries. Since our question asks the buying price of the lottery, it is possible that our measure of risk aversion is biased upward. However, if the bias is proportional to the reported price and is constant across individuals, the results should be unaffected.

aversion within the risk averse individuals which shows that preferences are very heterogenous with respect to risk.

It should be noted that this measure of risk requires no assumption on the form of the individual utility function and extends to risk-averse, risk-neutral and risk-loving individuals.[15] This lottery question has been used to study the relationship between risk aversion and several household decisions. Guiso and Paiella (2005), use the question on risk aversion to analyze occupation choice, portfolio selection, insurance demand, investment in education (in the linear OLS case) and migration decisions. They find substantial effects of this measure of risk aversion in ways that are consistent with the theory i.e. that more risk averse individuals choose lower returns in exchange for lower risk. They find for example that being risk averse increases the probability of being self-employed by 36% of the sample mean and the probability of holding risky assets by 42% of the sample mean. They also find that being risk averse as opposed to being risk neutral or risk prone (i.e. they use a risk-averse dummy), lowers education by one year on average. Guiso, Jappelli and Pistaferri (2002) show that risk aversion is negatively correlated with a measure of income risk (built from a question which asks about the expectations on future employment and income) i.e. risk averse individuals choose jobs with low income risk. Brunello (2002) estimates returns to schooling instrumenting schooling attainment with risk aversion under the hypothesis that risk aversion affects schooling costs but does not affect income if not through schooling. We will use the lottery question to explain schooling attainment and quantify the predictive power of risk aversion as opposed to other determinants of schooling.

Theoretically, the answer given by the individual may be partly affected by his/her time invariant degree of risk aversion but also partly affected by time varying differences in wealth/income endowment. Guiso and Paiella (2001) show that household income and wealth and individual characteristics have limited explanatory power.[16] Ultimately, they conclude that this measure of risk is a good proxy for the time invariant individual specific component of the attitude toward risk.

In order to verify this claim and obtain a measure of risk aversion which is orthogonal to earnings or wealth, we regress the response to the lottery

[15] It should also be noted that, given the answer to the lottery, it is possible to construct a measure of the Arrow-Pratt index of absolute risk aversion. This requires taking a second-order Taylor expansion of the relevant expected utility (Gollier, 2001).

[16] Interestingly, the main predictor of risk aversion is region of birth.

question, *bet*, on household wealth, household income and a dummy variable *Home_owner* which indicates whether the head is a home-owner.[17] The measure of household wealth *Hhold_wealth* includes all financial assets held by the household in 1995. The measure of household net income, *Hhold_income*, includes earnings, pensions and income from real and financial capital. The descriptive statistics of these variables are also in Table 1 in the Appendix. The results of the regression (not shown) indicate that the amount of the bet, *bet*, is positively related to household income and to household wealth and it is insignificantly related to home-ownership. However, the three variables account for only 2% of the total variance of *bet*. The residual of this regression is variable *bet_residual* in Table 1.

Guiso and Paiella (2001) discuss in details the main advantages of this estimate of absolute risk aversion relative to those already in the literature. They underline that the lottery represents a relatively large risk. In fact, ten million lire corresponds to just over 5,000 dollars and the ratio of the expected gain of the hypothetical lottery to the annual average Italian household consumption is 16 percent. This is considered an advantage since expected utility maximizers may behave as risk neutral individuals with respect to small risks even if they are risk-averse to larger risks. Thus, facing consumers with a relatively large lottery may be a good strategy to elicit risk attitudes.

Apart from the lottery question, we use information on the level of education attained by the head of household, as well as variables such as age, gender, region of birth, parental education and parental occupation. This set of variables is comparable to those which are used in US studies based on the National Longitudinal Survey (NLS). We select the sample of all heads with a valid answer to the lottery question (3,458) and eliminate those who report a missing value in any of the following variables: education, age, gender, region of birth, education and occupation of the head's father and mother. This selection process leaves us with a final sample of 3,288 heads of household.

The schooling variable takes values for 1 to 6 corresponding to no education, elementary school (typically attained at 11 years of age), junior high school (attained at 14), high school (attained at 18), university degree (attained at 23-24) and post-university degree.

Table 1 in the Appendix shows the descriptive statistics of the sample. In the estimation we use dummy variables derived from the original variables. There are six dummy variables - $edu1$ to $edu6$ - for the level of education at-

[17] This is also done in Brunello (2002).

tained by the individual (no title, elementary school, junior high school, high school, university degree, post-college degree), three dummies - *north*, *centre* and *south* - for the region of birth, one age dummy ($age45more = 1$ if age of head more than 45 in 1995) and one sex dummy ($female = 1$).[18] In addition we have one dummy each - edu_father and edu_mother - respectively for the level of education attained by the individual's father and mother (less than high school=0, high school or more=1), and four occupation dummies for blue collar, white collar, self employed and unoccupied for parents' occupation. These variables are denoted $bc\_father, wc\_father, se\_father, u\_father$ for the father and $bc\_mother, wc\_mother, se\_mother, u\_mother$ for the mother.

4 The Econometric Model

In this section, we present the econometric model. As schooling attainments are reported according to six (ordered) levels, we model schooling decisions with a reduced-form dynamic discrete choice model and we use a hazard function model of grade transition. The grade transition model admits a semi-structural interpretation and may be regarded as an approximation to a sequential dynamic optimization model. In recent work, Heckman and Navarro (2005) show that it is possible to conceive non-parametric (or semi-parametric) identification of reduced-form dynamic discrete choice models, such as discrete hazard functions, under certain conditions.[19]

Our strategy is to model the grade transition function as flexibly as possible. It is important to note that an alternative would be to model schooling as an ordered static discrete choice (say an ordered logit or an ordered probit) in which the error term may be viewed solely as cross sectional differences in tastes or abilities. However, in such an ordered model, the error term does not allow for randomness in the schooling decisions from the perspective of the agent and should be interpreted solely as randomness from the perspective of the researcher. The ordered probability model specification is usually

[18] The reason we introduce a dummy for the heads of household older than 45 in 1995 is that presumably they started college (if they ever attended it) before 1968. Before 1968 legal restrictions limited the accession to college only to those who had a high school degree in classical or scientific studies, since 1968 accession to college is open to any type of high school degree.

[19] Heckman and Navarro (2005) also show that similar identification results extend to structural dynamic programming models.

interpreted as a static (or myopic) model of schooling decisions.[20] For these reasons, we favor the hazard specification, although we also present some results obtained with the ordered logit.

Our method is based on two fundamental assumptions; namely that choices be made sequentially and that risk aversion be separable in a time invariant component (relevant when schooling decisions were made) and a possibly time variable component, which may reflect changes in permanent wealth.[21]

We see at least three main advantages to our approach.

First, it does not require to specify individual preferences but only requires that the measure of preference heterogeneity is a good proxy for the ordering of the persistent degree of risk aversion across individuals.

Secondly, it neither requires to model the distribution of labor market outcomes, nor to assume that the distribution of the labor market outcome variables, which are realized in the post-schooling periods, is actually known by the agents at the time of the schooling decisions.

Finally, we do not need to assume that the persistent unobserved (to the econometrician) heterogeneity term(s) affecting labor market outcomes belong(s) to the information set of the agent when schooling decisions were made. Our estimation strategy is therefore consistent with schooling decisions made under imperfect information about individual specific skills.[22]

The model allows for different types of factors; measured preference heterogeneity, family characteristics (parents' education and occupation), gender, regional effects, cohort effects and, finally, persistent individual unobserved heterogeneity.

With six (ordered) levels of schooling, we are able to estimate five different hazard rates. The conditional probability of stopping at grade g of individual i (a hazard rate), denoted H_g, is simply

$$H_{g,i} = \Lambda(U_{g,i}) \text{ for } g = 1, 2, ..5 \tag{3}$$

where

$$U_{g,i} = \alpha_{g,i} + \beta_g' X_i \tag{4}$$

[20] This is discussed in Cameron and Heckman (1998).

[21] However, in line with what Guiso and Paiella (2005) have shown, the measure of risk aversion available in the SHIW data appear to be mostly affected by the permanent part.

[22] This issue is analyzed formally in Cunha, Heckman and Navarro (2005).

The term $\alpha_{g,i}$ represents an individual/grade specific intercept term, X_i is a vector of observable characteristics, and β'_g represents a grade specific vector of parameters measuring the effects of these characteristics. We assume that

$$\alpha_{g,i} = \alpha_g + \theta_i$$

and that θ_i is drawn from an unknown distribution which is approximated by a discrete distribution with K points of support.[23] This approach amounts to the estimation of a vector of grade level specific intercept terms for each type. As we include an intercept term in the transition probability, we normalize one support point (namely θ_1) to 0.[24] In line with Heckman and Navarro (2005), we estimate $\Lambda(.)$ as flexibly as possible. As advocated by Geweke and Keane (2000), $\Lambda(.)$ is approximated with a mixture of 5 normal random variables with // free parameters: that is

$$\Lambda(.) = \sum_{m=1}^{M} P_m^* \cdot \Phi(\mu_m, \sigma_m)$$

where P_m^* is the mixing probability and $\Phi(\mu_m, \sigma_m)$ denotes the normal cumulative distribution function. To obtain identification, we impose that for on m, $\Phi(\mu_m, \sigma_m)$ denotes is the standard normal $\Phi(0, 1)$ since X$_i$ contains an intercept term and impose a labeling restriction.

Denoting the measured wealth and risk aversion (as of 1995) by $W_{i,95}$ and R$_{i,95}$ respectively, we assume that

$$W_{i,95} = X^w_{i,95}\beta^w + \varepsilon^w_{i,95} \tag{5}$$

and

$$R_{i,95} = X^r_{i,95}\beta^r + \varepsilon^r_{i,95} \tag{6}$$

[23] As typically found in most empirical applications dealing with a univariate duration endogenous variable, it has been found that $K = 2$ is sufficient to characterize unobserved heterogeneity.

[24] Obviously, the probability of transiting from one grade level (g) to the next ($g + 1$), the grade transition (continuation) probability, is simply

$$1 - H_{g,i} = \frac{1}{1 + \exp(U_{g,i})}$$

We sometimes refer to the grade transition probability as a "continuation" probability. The hazard rate is sometimes referred to as a "termination" probability.

ε_{it}^{w} and $\varepsilon_{i,95}^{r}$ are distributed with density $f^{w}(\varepsilon_{it}^{w}\ )$and $f^{r}(\varepsilon_{i,95}^{r})$. We approximate each density with a mixture of 5 unrestricted normal densities;

$$f^{s}(\varepsilon_{i,95}^{s}) = \sum_{m=1}^{M} P_{m}^{s} \cdot \phi(\mu_{m}, \sigma_{m}). \tag{7}$$

We estimate the model by maximum (mixed) likelihood techniques. Defining six different schooling indicators from the lowest schooling level (d_{1i}) to the highest (d_{6i}), the contribution to the likelihood for an individual i who has completed level g, is denoted L_i, and is equal to

$$L_i = \sum_{k=1}^{K} p_k \cdot [\Pi_{s=1}^{g-1}(1 - H_{s,i}(X_i, \theta_k))^{s} \cdot H_{g,i}(X_i, \theta_k)] \tag{8}$$

where θ_k is the type specific support point and where the type probability, p_k, is specified as $\frac{\exp(p_{0k})}{1+\exp(p_{0k})}$. Given the form of the hazard specification (equation 3), it is important to note that the sign of the parameter estimates indicates the direction of the effect of a variable on the exit rate out of school. So a negative estimate will typically imply a positive effect on expected grade completion, and in particular, on the probability of reaching higher education.

Defining g^h as the number of years of schooling required to attend higher (post-secondary) education and letting S_i as the number of years of schooling completed as of 1995, the probability of attending post-secondary education, denoted $\tilde{P}_i(X_i, \theta_i)$, is given by

$$\tilde{P}_i(X_i, \theta_i) = \Pr(S_i \geq g^h \mid X_i, \theta_i) = \Pi_{s=0}^{g^h}(1 - H_{s,i}(X_i, \theta_i)) \tag{9}$$

Given estimates of the type specific population proportions and support points (the p_k's and the θ_k's), the individual specific higher education attendance probability, $\tilde{P}_i(X_i, \theta_i)$, can easily be evaluated. To perform the variance decomposition, we choose the individual specific log odds ratio, $G(X_i, \theta_i)$, which is

$$G(X_i, \theta_i) = \left[\ln\left(\frac{\tilde{P}_i(X_i, \theta_i)}{1 - \tilde{P}_i(X_i, \theta_i)}\right)\right] \tag{10}$$

where $G(X_i, \theta_i)$ is treated as an unknown regression function to be estimated. In Section 6, we decompose the total variation in $G(X_i, \theta_i)$ into the different contributions of some specific group of variables; preference heterogeneity, family characteristics (parents' education and occupation), socioeconomic variables (gender, regional effects, cohort effects) and, finally, persistent individual unobserved ability heterogeneity. This approach allows us to obtain a ranking of the various groups of variables and, in particular, to establish if (and to what extent) differences in risk aversion are important.

5 An Overview of the Different Model Specifications

In order to obtain a clear picture of the effect of risk aversion on school attendance, we first estimate a simple version of the grade transition model with 5 grade specific intercept terms and with a common set of parameters assumed to be constant across all grade levels (where $\beta'_g = \beta \; \forall g$). We refer to it as Model 1. It was estimated first with the individual bet as the measure of risk aversion and then re-estimated with a discretized measure of the bet information (in 6 categories) so to minimize the impact of measurement error. The results are found in Table 3.

As a second step, we estimated a more flexible version of the model where the regressors have separated effects by grade level. The specification reported herein (in Table 4) allows for the effect of the regressors to change at the level prior to college enrolment. In other words, it allows for the marginal effect of risk aversion heterogeneity to change as an individual who has competed the 4th level is deciding to continue to the 5th level (corresponding to higher education).[25]

6 How Does Risk Aversion Affect Grade Progression?

In this section, we concentrate on the discussion of the parameter estimates of the grade transition model. Basically, and in line with the theoretical literature, we present evidence that schooling attainments decrease with risk aversion and we evaluate the robustness of the results to alternative measures of risk aversion and market skills. We report that the effect of risk aversion

[25] The estimations are performed using a FORTRAN program. However, they are easily doable using standard econometric softwares such as SAS or STATA.

is not magnified at higher grade levels. Finally, we stress that the range of higher education participation probabilities spanned by the 10th-90th percentile range of the risk aversion measure is ten to twenty times smaller than the corresponding difference between the two heterogeneity types.

6.1 Estimates from a Simple Model

The results obtained for the simplest model specification are found in Table 3. The gradual increase in the intercept terms (α_1 to α_5) from level 1 (-3.94) to level 5 (9.30) indicates that the schooling decision process is characterized by an increasing termination rate (given unobserved heterogeneity). The importance of unobserved heterogeneity is readily seen from the support point estimate for type 2 individuals (-4.52) along with the type 1 probability equal to 0.75. This implies that the population is clearly split between a high schooling attainment (low hazard rate) sub-population made of type 2 individuals, and a lower schooling attainment (higher hazard rate) sub-population made of type 1 individuals. The importance of unobserved persistent heterogeneity is a well known feature of most studies based on US data. A large number of studies set in a dynamic framework point out that permanent unobserved heterogeneity, which may represent unobservable factors such as individual specific taste for schooling, academic ability, motivation, or any other unobservable trait which is time-invariant, is indeed the major determinant.[26]

As documented in many empirical studies, grade termination is lower for those whose parents have achieved higher education. The parameter estimates for father's and mother's education are equal to -2.47 and -1.48 respectively. This positive correlation between individual schooling attainments and parents education is well established in simple correlation analysis (Kane, 1994), in reduced-form dynamic models such as Cameron and Heckman (1998, 2001) as well as in structural dynamic programming models such as Eckstein and Wolpin (1999) and Belzil and Hansen (2002). Grade continuation is also higher for those who have a parent who worked in a white collar occupation (the omitted category). This is readily seen upon looking at the positive effect (on the hazard rate) of the binary variables for all other occupation types. As expected, we find that individuals living in the North

[26] Cameron and Heckman (1998 and 2001), Eckstein and Wolpin (1999) and Belzil and Hansen (2002 and 2003).

(the most economically developed region of Italy), when compared to those who live in central regions, obtain more schooling. Finally, both females and younger cohorts appear to have lower grade termination rates. However, given the objectives of the paper, these estimates do not raise immediate interest and we do not discuss them in detail.

The parameter that raises most interest is the effect of preference heterogeneity as measured by the individual specific value attached to the lottery. Given unobserved heterogeneity and other measurable characteristics, we find that those individuals who attach a higher value to the lottery (those who are less risk averse) tend to have a lower grade termination rate. The estimate, equal to -0.565, is highly significant, and therefore indicates that more risk averse individual obtain less schooling. This is in agreement with conventional wisdom. However, we do not know of any comparable results, where the degree of schooling selectivity is directly tied to an observable measure of risk aversion, in the empirical literature.[27]

It is also interesting to note that our results appear to conflict with those reported in a somewhat related literature that uses smoking behavior as an instrument for schooling, in order to estimate the return to schooling.[28] The first stage regressions often indicate that schooling is inversely related to smoking behavior and this finding could be interpreted as evidence that risk averse individuals (those who smoke less) obtain more schooling. However, it should be noted that smoking is an endogenous variable, which is likely to be affected by several factors including intrinsic taste for smoking, parents' background (including education), teenage schooling attainments (performance in school) and other individual specific factors such as risk aversion and discount rates. And indeed this literature itself associates smoking with higher discount rates rather than with pure risk aversion. It is therefore not certain that changes in schooling induced by smoking differences are solely due to risk aversion and, therefore, that standard assumptions (such as monotonicity and homogeneity) made in the IV literature would be valid in this context.

[27] For instance, in the theoretical model of Lehvari and Weiss, this relationship would be derived from differentiating the expected utility of staying in school with respect to a measure of concavity of the utility function. In Belzil and Hansen (2004), the effect of risk aversion on schooling is obtained upon differentiating the school attendance probability (involving a closed-form solution to the value function) with respect to a parameter representing absolute (or relative) risk aversion.

[28] Fersterer and Winter-Ebmer (2003), Chevalier and Walker (1999) and Evans and Montgomery (1994).

Our results, on the other hand, illustrate a marginal effect of risk aversion, holding all other factors constant. They are certainly not incompatible with the hypothesis that young individuals coming from poorer background and less educated families tend to smoke more (given a fixed degree of risk aversion).

At this stage, the non-linearity of the model is preventing to see the clear effects that both the risk aversion measure and other attributes may have on schooling transition probability. The analysis of these related marginal effects are delayed to Section 7.

6.2 Some Robustness Analysis

In order to minimize the impact of potential measurement error, we estimated the same model with a discrete measure of risk aversion. We split the sample into 6 groups according to their measure of risk aversion. The groups are 0 to 150 (group 1), 151 to 500 (group 2), 501 to 1000 (group 3), 1001 to 3000 (group 4), 3001 to 5000 (group 5) and those reporting a bet above 5000 (group 6). This will allow us to perform an analysis that may be less dependent on outliers (extremely risk averse or extremely risk loving individuals). The results are found in the right-hand side of Table 3.

The results are consistent with those reported before. The range of estimates from Group 2 to Group 6 (-0.2559 to -1.2061) also indicate that schooling termination rates increase with risk aversion. A informal look at the standard errors reveal that, except for group 3 and 4, most contiguous groups are significantly different from each other.

We also considered an alternative measure of the monetary value associated to the lottery described above. In order to take into account that individual wealth and income may affect substantially the degree of risk aversion measured by the answer to the lottery, we re-estimated the first model specification of Table 3. If the monetary value is truly affected by wealth, then a measure of risk aversion purged of wealth effects may turn out to be closer to a truly individual specific (time invariant) measure of risk aversion. However, using the residual of the regression of the monetary value of the individual bet on various measures of wealth will change the scaling of the variables.

The estimates, found in Table 10, indicate clearly that the results discussed earlier are indeed robust to the introduction of this alternative measure of risk aversion. The effect of the individual specific monetary value is equal to -0.28 and implies again that risk aversion is negatively correlated with schooling attainment. Most of the remaining parameters have practically not changed.

6.3 Does the Effect of Risk Aversion Change with Grade Level?

We now examine the estimates obtained from the more flexible specifications (Model 2) found in Table 4. It allows for a common set of parameters from level 1 to level 3 and a different set common at level 4 and level 5. This specification enables the effect of risk aversion to change at the grade level where the strategic decision to enter higher education is made. Ultimately, we use the implied marginal effects and the variance decompositions to establish a ranking between the group of variables in terms of their importance in explaining schooling attainments. As will become clear later, we show that, despite a statistically significant effect, differences in attitudes toward risk are virtually unimportant. Unobserved persistent factors and family human capital play a substantially larger role.[29]

The estimates of Table 4 suggest that risk aversion decreases grade continuation both at the low and the higher grade levels. The estimates, -0.824 (level 1 to 3) and -0.141 (level 4 and beyond), are both significant at any confidence level. In order to fix ideas, it is useful to compute average grade termination (hazard) rates over the relevant range of the risk aversion heterogeneity variable. To get a clear picture, we report the average hazard rates at the 10th, the 50th and the 90th percentiles of the risk aversion variable. These estimates are found in the left-hand side of Table 5. They illustrate clearly the weak effect of risk aversion on grade termination. As an example, the average probability of terminating at grade level 4 fluctuates between 0.883 for someone endowed with a rather extremely low value for the risk aversion indicator (a bet which is ranked at the 10th percentile) and 0.878 for someone at the 90th percentile. The equivalent ranges are slightly higher for level 3 and level 2, but they remain small. Overall, there is no evidence

[29] We have also estimated a more general model with the effect of the regressors changing at each possible grade level. The basic results remaining the same, and for transparency purposes, we decided to report the results of the specification involving a smaller number of parameters.

that the effect of risk aversion is magnified as one approaches the decision to enter higher education.

For the model where the risk aversion variable is split in 6 groups, we also still find a negative effect of risk aversion on school completion. However, the group-specific parameter estimates are much more precisely estimated at the lower levels (1, 2 and 3) than at the higher level. The range of estimates for the low level is between -0.15 for group 2 and -1.44 for group 6. The ordering is practically monotonic (except for group 3 and group 4 who are practically equal). At the higher level (level 4 and 5), all estimates are negative but the values are erratic. Some parameter estimates for (group 2 and group 6) are not even significantly different from group 1 (the benchmark group).

For this model, we report hazard rates for group 1, 3 and 6 (Table 6). This may not be perfectly comparable with the 10th-90th percentile range, but it is still a good indication. Consistent with what is reported in Table 5, the maximum range in hazard rates is around 0.12 at level 2 (0.29 for group 1 and 0.17 for group 6). The range between group 1 and group 6 is 0.03 at level 4. Overall, differences in risk aversion appear to translate into modest differences in school continuation probabilities.

6.4 How do Differences in Risk Aversion Compare with Unobserved Differences in Ability?

For a sake of comparison, we perform a similar exercise with the distribution of unobserved heterogeneity. As our estimation procedure splits the population into two types according to unobserved heterogeneity, it is easy to compute an average hazard rate for each type of individuals, and for each grade level. This will allow us to obtain a relative measure of the importance of preference heterogeneity as opposed to ability heterogeneity.

The type specific grade termination rates are found in Table 5 and Table 6. The difference in probability of stopping school between type 1 and type 2 exceed the difference recorded between the 10th and 90th percentiles of the risk aversion measure at all grade levels. In the case where the actual bet is used (Table 5), the difference in probability of stopping at grade level 4 (just before entering higher education) between type 1 and type 2 individuals is around 15 times larger than for the 10th/90th percentile difference. With the discretized value, the type1/type 2 difference is around 10 times larger than the difference between group 1 and group 6.

Schooling decisions appear to be overwhelmingly dominated by skill differences as opposed to differences in attitudes toward risk. At this stage, there is clear evidence that schooling attainments are much more affected by differences in ability than by differences in attitude toward risk.

7 What Fraction of Schooling Attainments is Explained by Risk Aversion?

In order to obtain a more global picture of the importance of risk aversion, we now proceed with a precise variance decomposition. Using the estimates of both Table 3 and Table 4, we compute individual specific probabilities of achieving higher education and transform them into the logarithm of their odd's ratios. These probabilities differ across individuals according to the various groups of variables. Using regression techniques, it is easy to compute how much of the individual differences in odd's ratios are explained by each component separately. In what follows, we show that risk aversion is by far the least important of the group of factors on which we focussed our analysis.

The results are found in Table 7 (for Model 1) and Table 8 (for Model 2). With the actual bet used as a regressor in Model 1, we find that unobserved persistent differences account for 60% of the cross-sectional differences in the probability of entering higher education (Table 7) and 72% in Model 2 (Table 8). With the discretized value, unobserved heterogeneity remains the most important variable and the corresponding values are 55% (Table 7) and 72% (Table 8).

Consistent with the very small marginal effects already noticed, the portions of the total variances explained by differences in risk aversion are very small, they vary between 2% (model 1/with the actual bet) and 7% (model 2/discretized value).[30]

Interestingly, parental human capital accounts between 20% and 30% of the explained differences. It is, by far, the second most important group of factors. The ranking appears quite robust. Aside from the residual socioeconomic factors (sex, region and cohort) which account for 3% to 7%, risk aversion is the least important factor among those on which we focussed our

[30] One alternative method would be to compute a regression on all components and evaluate the loss in explanantory power when each group of variable is removed individually. This method leads to an identical ranking.

analysis.

8 Concluding Remarks

In this paper, we present evidence that schooling attainments decrease with risk aversion. This result is in line with the early theoretical literature. However, we show that, despite a statistically significant effect, differences in attitudes toward risk are not that important. Unobserved persistent factors, market skills and family human capital play a substantially larger role.

While interesting, these answers deserve some interpretation and also raise one fundamental question; Why is the level of risk associated to schooling, as perceived by individuals, **so small**?

One possible answer is that despite the intrinsic risk faced by those who decide to enter higher education, workers may have the perception that schooling reduces wage (or earnings) dispersion. In other words, young individuals regard schooling as an insurance and the marginal risk associated to higher grade enrollment is small. If this is true, it would be interesting to see if this is specific to Italy only or if this may extend to other countries.

There is another possible answer. It is conceivable that entering higher education may preserve the option value of choosing occupations, sectors or jobs (firms) which are characterized by stable and safe earnings profiles. In other words, the relevant decisions that involve differences in attitudes toward risk are occupation and/or sectoral employment choices. If these choices are made posterior to the decision to enter higher education, schooling decisions, as such, will not reveal selectivity based on differences in risk aversion.

While we believe that the analysis presented in this paper is interesting in its own right, we recognize that answering these questions would be important. However, it would require a more sophisticated analysis and access to similar data from other countries. This may be an interesting, but challenging, avenue for future research.

References

[1] Baker, M. and Angelo Melino (2000) "Duration Dependence and Nonparametric Heterogeneity: a Monte Carlo Study" *Journal of Econometrics*, 96(2), 357-393.

[2] Belzil, Christian (2005) "The Return to Schooling in the Structural Literature: A Survey" forthcoming in *The European Economic Review.*

[3] Belzil, Christian and Hansen, Jörgen (2004) "Risk Aversion, Education and Earnings Dispersion" *Research in Labor Economics, volume 23.*

[4] Belzil, C. and Jörgen Hansen (2003) "Structural Estimates of the Intergenerational Education Correlation" *Journal of Applied Econometrics,* 8, 679-696.

[5] Belzil, C. and Jörgen Hansen (2002) "Unobserved Ability and the Return to Schooling" *Econometrica,* 70(5), 2075-2091.

[6] Ben-Porath, Yoram (1967) "The Production of Human Capital and the Life Cycle of Earnings" *Journal of Political Economy,* 75(4), 352-365.

[7] Brunello, G. (2002) "Absolute Risk Aversion and the Returns to Education" *The Economics of Education Review*, 21, 635-640.

[8] Cameron, S. and James Heckman (1998) "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males" *Journal of Political Economy,* 106 (2), 262-333.

[9] Cameron, S. and James Heckman (2001) "The Dynamics of Educational Attainment of Black, Hispanic and White Males" *Journal of Political Economy*, 109(5), 455-499.

[10] Cameron, S. and Chris Taber (2001) "Estimation of Educational Borrowing Constraints using Returns to Schooling" Working Paper, Northwestern University.

[11] Chay, K. and David Lee (2000) "Changes in Relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials", *Journal of Econometrics*, 99, 1-38.

[12] Chen, Stacey H. (2003) "Estimating Wage Volatilities for College versus High School Careers" Working paper, SUNY Albany.

[13] Chevalier, A. and Ian Walker (1999). "Further Results on the Returns to Education in the UK" mimeo, University of Warwick.

[14] Chiswick, B.R. and Jacob Mincer (1972) “Time-series changes in Personal income inequality in the United states from 1939, with Projection 1985" *Journal of Political Economy,* 80(3), S34-S66.

[15] Cunha, F., J. Heckman and Salvador Navarro (2005) ”Separating Uncertainty from Heterogeneity in Life Cycle Earnings” NBER Working Paper 11024.

[16] Davis, S. and David Willen (2000) “Risky Labor Income and Portfolio Choice” Working Paper, University of Chicago

[17] Eckstein, Zvi and Kenneth Wolpin (1999) “Why Youths Drop Out of High School: the Impact of Preferences, Opportunities and Abilities” *Econometrica*, 67(6), 1295-1339.

[18] Evans, W.N.and E. Montgomery (1994) ”Education and Health: Where There’s Smoke There’s an Instrument” NBER WP 4949.

[19] Fersterer, J. and Rudolf Winter-Ebmer (2003) ”Smoking, Discount Rates, and Returns to Education” Economics of Education Review, 22(6), 561-566.

[20] Flinn, Christopher (2002) ”Labour Market Structure and Inequality: A Comparison of Italy and the U.S.” *Review of Economic Studies,* 69(3), 611-645.

[21] Geweke, J. and Michael Keane (1995) “Bayesian Inference for Dynamic Discrete Choice Models without the Need for Dynamic Programming" Working Paper, Federal reserve Bank of Minneapolis.

[22] Gollier, Christian (2001), *The Economics of Risk and Time*, The MIT Press.

[23] Gould, E., Moav, O. and Bruce Weinberg (2001) “Precautionary Demand for Education, Inequality and Technological Progress" *Journal of Economic Growth,* 6, 285-315.

[24] Guiso, L. and Monica Paiella (2001) ”Risk Aversion, Wealth and Background Risk” CEPR Discussion Paper 2728.

[25] Guiso, L. and Monica Paiella (2005) ”The Role of Risk Aversion in Predicting Individual Behaviour” Banca d'Italia Temi di Discussione 546.

[26] Guiso, L., T. Jappelli and L. Pistaferri (2002) ”An Empirical Analysis of Earnings and Employment Risk” Journal of Business and Economic Statistics, 20, 1-13.

[27] Ham, J. and Robert Lalonde (1996) “The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training” *Econometrica*, 64(1), 175-205.

[28] Harmon, C., Hogan, V. and Ian Walker (2003) ”Dispersion in the Economic Returns to Schooling” *Labour Economics*, 10(2), 205-214.

[29] Heckman, J., Lochner, L. and Petra Todd (2005) “Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond" NBER Working Paper 11544.

[30] Heckman, J. and Salvador Navarro (2005) “Dynamic Discrete Choices and Dynamic Treatment Effects" forthcoming in *Journal of Econometrics.*

[31] Heckman, J. and Ed Vytlacil (1998) “Instrumental Variables Methods for the Correlated Random Coefficient Model” *Journal of Human Resources*, 33 (4), 974-987.

[32] ISTAT (2003) ”Lo Stato dell'Universita”' Istituto Nazionale di Statistica, Roma.

[33] Johnson, Thomas (1978) “Time in School: The Case of the Prudent Patron” *American Economic Review,* 68 (5), 862-872.

[34] Kagel, J. H. and Alvin E. Roth (eds.) (1995) ”The Handbook of Experimental Economics” Princeton University Press, Princeton.

[35] Kocherlakota, Narayana (1996) ”The Equity Premium: It's Still a Puzzle” *Journal of Economic Literature,* 24 (1), 42-71.

[36] Kodde, David (1986) “Uncertainty and the Demand for Education” *The Review of Economics and Statistics,* 68(3) 460-467.

[37] Hartog, J., van Ophem, H. and Simona Bajdechi (2004) “How Risky is investment in Human Capital” Working Paper, Tinbergen Institute.

[38] Lehvari, D. and Yoram Weiss (1974) “The Effect of Risk on the Investment in Human Capital” *American Economic Review,* 64(6), 950-963.

[39] Lemieux, Thomas (2005) "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skills?" forthcoming in *American Economic Review.*

[40] Mincer, Jakob (1974), *Schooling, Experience and Earnings,* New-York, Columbia University Press.

[41] Olson, L., White, H. and H.M. Sheffrin (1979) “Optimal Investment in schooling when Incomes are Risky” *Journal of Political Economy,* 87(3), 522-539.

[42] Palacios-Huerta, Ignacio (2003) “An Empirical analysis of the Properties of Human Capital Returns” *American Economic Review,* 93 (3), 948-964.

[43] Van den Berg, Gerard (2001) "Duration Models: Specification, Identification and Multiple Durations" in J. J. Heckman and E. Leamer eds. Handbook of Econometrics, Volume V, North Holland, Amsterdam.

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
bet	3288	2513.083	4798.066	1	100000
bet_residual	3288	2.370	4745.169	-11421.100	98108.800
edu1	3288	0.006	0.078	0	1
edu2	3288	0.103	0.304	0	1
edu3	3288	0.320	0.467	0	1
edu4	3288	0.305	0.460	0	1
edu5	3288	0.238	0.426	0	1
edu6	3288	0.028	0.165	0	1
edu_father	3288	0.116	0.321	0	1
edu_mother	3288	0.073	0.260	0	1
north	3288	0.388	0.487	0	1
centre	3288	0.168	0.374	0	1
south	3288	0.443	0.497	0	1
female	3288	0.198	0.399	0	1
bc_father	3288	0.398	0.490	0	1
wc_father	3288	0.207	0.405	0	1
se_father	3288	0.361	0.480	0	1
u_father	3288	0.033	0.180	0	1
bc_mother	3288	0.089	0.285	0	1
wc_mother	3288	0.045	0.209	0	1
se_mother	3288	0.094	0.291	0	1
u_mother	3288	0.771	0.420	0	1
age45more	3288	0.609	0.488	0	1
Hhold_income	3288	48776.650	36472.330	-68000	771077.100
Hhold_wealth	3288	252549	383045.700	-139000	6785000
Home_owner	3288	0.634	0.482	0	1

Table 2: The individual specific value attached to the lottery: The distribution of *bet*

Deciles	*bet* (1,000 liras)
1	50
2	100
3	500
4	1000
5	1000
6	2000
7	3000
8	5000
9	5000

Table 3: Results from a grade transition model: Model 1

	Model 1 Continuous *bet*		**Model 1** Discretized *bet*	
	Coeff.	SE	Coeff.	SE
α_1	-3.9430	0.227	-8.4712	0.3054
α_2	-1.1173	0.2003	-5.6027	0.2307
α_3	0.8272	0.2293	-3.7080	0.2271
α_4	4.9246	0.3363	0.2429	0.1332
α_5	9.3026	0.5414	4.5238	0.3039
bet	-0.5648	0.1021	-	-
group 2	-	-	-0.2559	0.1049
group 3	-	-	-0.6926	0.0855
group 4	-	-	-0.7589	0.0938
group 5	-	-	-0.9583	0.0897
group 6	-	-	-1.2061	0.1484
edu_father	-2.4693	0.1857	-2.3747	0.2445
edu_mother	-1.4820	0.2259	-1.4882	0.2837
bc_father	0.7438	0.0564	0.7031	0.0634
bc_mother	1.3851	0.2238	1.3454	0.3671
se_father	0.066	0.0827	0.0492	0.1241
se_mother	0.911	0.1989	0.9254	0.3713
u_father	0.1319	0.2033	0.1281	0.332
u_mother	0.6952	0.1739	0.6564	0.3127
north	-0.2336	0.0669	-0.1896	0.1861
south	0.2775	0.0661	0.2492	0.1921
female	-0.5784	0.0882	-0.5609	0.1728
age45more	0.7439	0.0564	0.7063	0.0631
θ_2	-4.5168	0.2871	4.3337	0.3355
prob (θ_1)	0.7500	0.0093	0.3433	0.0908
Mean log lik.	-1.29017		-1.28210	
Sample size	3288		3288	

Notes: prob(θ_1) is the type-specific population proportion and θ_2 is the support point ($\theta_1$is normalized to 0). The groups are 0 to 150 (group 1, omitted), 151 to 500 (group 2), 501 to 1000 (group 3), 1001 to 3000 (group 4), 3001 to 5000 (group 5) and those reporting a bet above 5000 (group 6).

Table 4: Results from a grade transition model: Model 2

	Model 2 Continuous *bet*		**Model 2** Discretized *bet*	
	Coeff.	SE	Coeff.	SE
α_1	-5.4437	0.2109	-5.1124	0.4662
α_2	-2.2365	0.1209	-1.7957	0.4514
α_3	0.572	0.1351	1.0406	0.6324
α_4	5.8253	0.2999	6.2542	0.5216
α_5	9.8162	0.446	10.2988	0.6494
Level 1-3				
bet	-0.8240	0.0773	-	-
group 2	-	-	-0.1475	0.9416
group 3	-	-	-0.8331	0.2113
group 4	-	-	-0.8099	0.9207
group 5	-	-	-1.0987	
group 6	-	-	-1.4409	
edu_father	-3.2014	0.253	-3.1933	0.3428
edu_mother	-1.7012	0.279	-1.7691	0.4226
bc_father	1.0775	0.0661	1.0862	0.0838
bc_mother	1.8997	0.1705	1.8787	0.4805
se_father	0.0396	0.055	0.0219	0.1371
se_mother	1.4765	0.1782	0.5178	0.492
u_father	0.1405	0.1458	0.2487	0.4127
u_mother	1.1599	0.1278	1.1586	0.4348
north	-0.2245	0.0353	-0.1899	0.1832
south	0.4855	0.0568	0.4423	0.1671
female	-0.7214	0.0825	-0.6830	0.2196
age45more	1.0778	0.0631	1.0863	0.0821

Table 4: Continued

Level 4-5				
bet	-0.1412	0.0611	-	-
group 2	-	-	-0.3709	0.3002
group 3	-	-	-0.5877	0.2838
group 4	-	-	-0.5316	0.2313
group 5	-	-	-0.4784	0.2272
group 6	-	-	-0.2929	0.2687
edu_father	-2.0924	0.3149	-2.0713	0.2896
edu_mother	-1.5189	0.281	-1.5528	0.2680
bc_father	0.1051	0.0479	0.0987	0.0514
bc_mother	1.5377	0.3465	1.5351	0.2794
se_father	0.0256	0.0792	0.0334	0.0683
se_mother	0.5941	0.204	0.6094	0.2043
u_father	-0.036	0.242	0.0127	0.2159
u_mother	0.45	0.1351	0.4632	0.1541
north	-0.67	0.118	-0.7069	0.2892
south	-0.5616	0.1131	-0.5817	0.2969
female	-0.1776	0.0799	-0.1866	0.1136
age45more	0.1053	0.0449	0.0963	0.0553
θ_2	-4.5031	0.2141	-4.5350	0.3685
prob (θ_1)	0.4671	0.0936	0.6125	0.0924
Mean log lik.	-1.27404		-1.26646	
Sample	3288		3288	

Notes: prob(θ_1) is the type-specific population proportion and θ_2 is the support point ($\theta_1$is normalized to 0). The groups are 0 to 150 (group 1,omitted), 151 to 500 (group 2), 501 to 1000 (group 3), 1001 to 3000 (group 4), 3001 to 5000 (group 5) and those reporting a bet above 5000 (group 6).

Table 5: The effect of unobserved heterogeneity and risk aversion on grade hazard rate by level in Model 1-Continuous Values of *bet*.

Average grade transition probability by grade level for different values of risk aversion and by heterogeneity type (standard errors in parenthesis)

	Risk Aversion Percentile			Unobs. heterogeneity	
	10th	**50th**	**90th**	**Type 1**	**Type 2**
level 1	0.0221	0.0206	0.0151	0.0317	0.0004
	(0.012)	(0.012)	(0.012)	(0.07)	(0.07)
level 2	0.2388	0.2297	0.1927	0.3555	0.0090
	(0.04)	(0.05)	(0.05)	(0.04)	(0.08)
level 3	0.5430	0.5359	0.5052	0.7872	0.1180
	(0.07)	(0.08)	(0.08)	(0.12)	(0.10)
level 4	0.8830	0.8821	0.8780	0.8804	0.7252
	(0.07)	(0.08)	(0.08)	(0.11)	(0.11)
level 5	0.9928	0.9927	0.9923	0.9997	0.9801
	(0.08)	(0.05)	(0.06)	(0.07)	(0.10)

Note: The standard errors are calculated using the delta method.

Table 6: The effect of unobserved heterogeneity and risk aversion on grade hazard rate by level in Model 1-Discretized Values of *bet*.

Average grade transition probability by grade level for different values of risk aversion and by heterogeneity type (standard errors in parenthesis)

	Risk Aversion Group			Unobs. heterogeneity	
	group 1	**group 3**	**group 6**	**Type 1**	**Type 2**
level 1	0.0301	0.0139	0.0107	0.0313	0.0004
	(0.012)	(0.012)	(0.012)		
level 2	0.2902	0.1942	0.1664	0.3587	0.0098
	(0.04)	(0.05)	(0.05)		
level 3	0.5815	0.5048	0.4798	0.7892	0.1235
	(0.07)	(0.08)	(0.08)		
level 4	0.9036	0.8632	0.8716	0.9873	0.7062
	(0.07)	(0.08)	(0.08)		
level 5	0.9947	0.9912	0.9920	0.9997	0.9805
	(0.02)	(0.02)	(0.02)		

Note: The standard errors are calculated using the delta method.

Table 7: Variance decomposition of the odd's ratio of the conditional probability of entering higher education: Model 1

	Explanatory power	
Variables	**Model 1** Continuous *bet*	**Model 1** Discretized *bet*
Risk aversion	3%	2%
Ability heterogeneity	60%	72%
Parental human capital	28%	19%
Other socio/geographic	7%	3%

Note: The explanatory power refers to the R-square of the regression of simulated log odds ratio on each variable (or group of variables) taken individually.

Table 8: Variance decomposition of the odd's ratio of the conditional probability of entering higher education: Model 2

	Explanatory power	
Variables	**Model 2** Continuous *bet*	**Model 2** Discretized *bet*
Risk aversion	7%	4%
Ability heterogeneity	55%	71%
Parental human capital	31%	19%
Other socio/geographic	7%	3%

Note: The explanatory power refers to the R square of the regression of simulated log odds ratio on each variable (or group of variables) taken individually.

Table 9: Ordered Logit

	Coeff.	SE
α_1	-5.4990	0.2325
α_2	-2.7602	0.2072
α_3	-1.1823	0.2028
α_4	1.1357	0.2055
α_5	4.4945	0.3014
bet	4.22E-05	7.88E-06
edu_father	1.1761	0.1399
edu_mother	1.0531	0.1703
bc_father	-1.2595	0.1050
bc_mother	-0.5544	0.2101
se_father	-0.8889	0.1038
se_mother	-0.4655	0.2106
u_father	-1.0758	0.1960
u_mother	-0.1465	0.1800
north	0.1847	0.0927
south	-0.2791	0.0920
female	-0.4698	0.0828
age45more	-1.0225	0.0681
Log Likelihood	-4192.1179	
Sample	3288	

Notes: The dependent variable takes values 1 to 6, 1= no formal education and 6=post-college degree.

Table 10: Robustness Analysis: Results from a grade transition model: Model 1

	Model 1	
	Coeff.	SE
α_1	-8.9172	0.4441
α_2	-6.0891	0.4403
α_3	-4.1271	0.4305
α_4	-0.0105	0.3623
α_5	4.3576	0.4521
bet_residual	-0.2826	0.0757
edu_father	-2.4973	0.1869
edu_mother	-1.5336	0.2405
bc_father	1.5039	0.1093
bc_mother	1.3802	0.2613
se_father	0.0657	0.1279
se_mother	0.8862	0.2506
u_father	0.1199	0.2157
u_mother	0.688	0.2559
north	-0.2442	0.088
south	0.3007	0.0867
female	-0.6033	0.0923
age45more	0	0.1545
θ_2	4.5477	0.2758
prob(θ_1)	0.4623	0.1054
Mean log lik.	-1.29381	
Sample size	3288	

Notes: In this table we use *bet_residual*, the residuals of the risk variable *bet* on household income $Hhold\_income$, household wealth, $Hhold\_wealth$, and an indicator of home-ownership $Home\_owner$.